Pure Appl. Geophys.

DOI 10.1007/s00024-011-0327-x

Pure and Applied Geophysics

Forecast of Low Visibility and Fog from NCEP: Current Status and Efforts

Binbin Zhou,[1] Jun Du,[1] Ismail Gultepe,[2] and Geoff DiMego[1]

Abstract—Based on the visibility analysis data during November 2009 through April 2010 over North America from the Aviation Digital Database Service (ADDS), the performance of low visibility/fog predictions from the current operational 12 km-NAM, 13 km-RUC and 32 km-WRF-NMM models at the National Centers for Environmental Prediction (NCEP) was evaluated. The evaluation shows that the performance of the low visibility/fog forecasts from these models is still poor in comparison to those of precipitation forecasts from the same models. In order to improve the skill of the low visibility/fog prediction, three efforts have been made at NCEP, including application of a rule-based fog detection scheme, extension of the NCEP Short Range Ensemble Forecast System (SREF) to fog ensemble probabilistic forecasts, and a combination of these two applications. How to apply these techniques in fog prediction is described and evaluated with the same visibility analysis data over the same period of time. The evaluation results demonstrate that using the multi-rule-based fog detection scheme significantly improves the fog forecast skill for all three models relative to visibility-diagnosed fog prediction, and with a combination of both rule-based fog detection and the ensemble technique, the performance skill of fog forecasting can be further raised.

1. *Introduction*

Fog is infrequent but it may be a very hazardous weather condition related to all forms of traffic and on health. Central guidance from the National Centers for Environmental Prediction (NCEP) on fog thresholds is being considered and particularly emphasized by the National Weather Service (NWS) of the National Oceanic and Atmospheric Administration (NOAA), and in NextGen (Souders, 2010), a future Air Traffic Management System of the Federal Aviation Administration (FAA), in the United States. However, fog is still not a part of the documents of NCEP central guidance, due to its complexity and limitation of computational resources. Instead, it is only diagnosed locally by forecasters either through subjective visibility forecasts or through other variables from model output such as MOS (Model Output Statistics). Nevertheless, effort to add it to NCEP's central guidance is considered to be important. As a step forward in response to the request from NWS and NextGen, low visibility/fog forecast has been experimentally implemented, tested and validated using NCEP operational models.

Currently, the visibility-liquid water content (LWC) relationship of (Stoelinga and Warner, 1999) is used in horizontal visibility computations in all the NCEP models. However, studies have shown that this visibility computation has large errors, particularly in the situation of fog when droplet number concentration (N_d) is not considered (Gultepe and Isaac, 2004; Gultepe *et al.*, 2006). Besides the error from the visibility computation, a bias in the model LWC near the surface is another source of errors. The visibility computation error can be reduced by applying Gultepe's visibility versus LWC and N_d parameterization (Gultepe *et al.*, 2006, 2009), whereas the reduction of model LWC error is extremely difficult due to a lack of fog physics for all fog types, model bias and low resolution of the operational models. To overcome these drawbacks, we have recently developed a rule-based fog detection scheme (Zhou and Du, 2010). The rule-based fog detection scheme is a combination of rules related to surface LWC, relative humidity with respect to water (RH), wind speed, and fog top (Zt) and base (Zb) heights for various fog types.

The second improvement effort is extending the NCEP Short Range Ensemble Forecast System (SREF, Du *et al.*, 2006) to fog forecasting. Because

[1] IMSG at EMC/NCEP/NWS/NOAA, 5200 Auth Road, Camp Springs, MD 20746, USA. E-mail: Binbin.Zhou@noaa.gov
[2] Cloud Physics and SWRS, MRD, Environment Canada, Toronto, ON, Canada.

Published online: 19 May 2011

of the chaotic and highly nonlinear nature of the atmospheric system, initially small differences in either initial conditions (ICs) or the model itself can amplify over forecast time and become large after a certain period of time. Since intrinsic uncertainties always exist in both ICs and model physics, a forecast by a single model run always has uncertainties too. Such forecast uncertainties vary from time to time, from location to location, and from case to case. A dynamical way to quantify such flow-dependent forecast uncertainties is to use an ensemble forecasting (Leith, 1974). Instead of using a single model simulation, multiple model integrations are performed that were initiated with slightly different ICs and/or based on different model configurations. Given the intrinsic uncertainties in model forecasts, and the fact that fog forecasting is believed to be extremely sensitive to the initial conditions and the physics schemes used in a prediction system (Bergot and Guedalia, 1994; Bergot, 2005), it is strongly desirable to have fog prediction as a part of the NCEP's ensemble framework.

Both the rule-based fog detection scheme and the ensemble application have been tested and evaluated in the World Meteorological Organization (WMO)'s 2008 Beijing Olympic Game Research and Demonstration Project (B08RDP) over China with 7 months of data from 13 Chinese cities (Zhou and Du, 2010). The evaluations have shown that the rule-based fog detection scheme could improve the fog forecasting score by a factor of two while its combination with the ensemble technique could add extra value to fog predictions. After B08RDP was finished, these two techniques, individually and combined, were further tested and evaluated over North America for NCEP's regional models and the ensemble forecast system.

The objective of this paper is to evaluate (1) performance of fog prediction using the visibility-LWC relationship from the current NCEP's regional models, (2) rule-based fog detection scheme, and (3) ensemble forecast technique for fog detection over the North American domain. This paper is organized as follows: section 2 is for the configurations of the models and the ensemble forecast system involved, section 3 is for the evaluation method for the results, and section 4 is for the results and discussion, followed by the conclusion section.

2. *Configurations and Methods*

2.1. *Configuration of Regional Models*

The three regional operational models over North American domain (see Fig. 1) used for low visibility (Vis) and fog prediction verifications are (1) 13-km resolution Rapid Update Cycle model, or RUC-13 (Benjamin, 2003), (2) 12-km resolution North American Mesoscale model, or NAM-12 (Rogers *et al*., 2005), and (3) 32-km resolution Non-hydrostatistic Mesoscale Model, or NMM-32 (Janjić, 2001). RUC-13 runs hourly, specifically for aviation weather forecasts. NAM-12 runs four times (00, 06, 12 and 18Z) per day to provide central guidance to regular weather for all local forecasters in the United States. NMM-32 is NCEP's WRF (Weather Research and Forecasting) model, which is also one of the base models in the SREF system running four times (03, 09, 15 and 21Z) per day to generate both regular and aviation weather guidance. In fact, NAM-12 is also a WRF-NMM based model but runs in different horizontal resolution from NMM-32. The parameterization schemes employed in both NAM-12 and RUC-13 are listed in Table 1.

2.2. *Visibility Computation Method*

Currently, there is no direct fog prediction algorithm used in either the NCEP's regional models or SREF system. Instead, the visibility computation in these models is based on the algorithm of (Stoelinga and Warner, 1999) that uses rain, snow, and cloud water or ice amount. In the case of fog, the Kunkel (1984) Vis-LWC parameterization is employed, where the LWC is the value at the lowest model level.

2.3. *Rule-Based Fog Detection Scheme*

Although one hopes that LWC at the lowest model level can be explicitly used for fog calculation, experience tells us that the visibility-LWC approach doesn't work well in the operational models mainly for two reasons: one is the too coarse model spatial resolution and the other is the lack of sophisticated fog physics. As a result, LWC from the models is

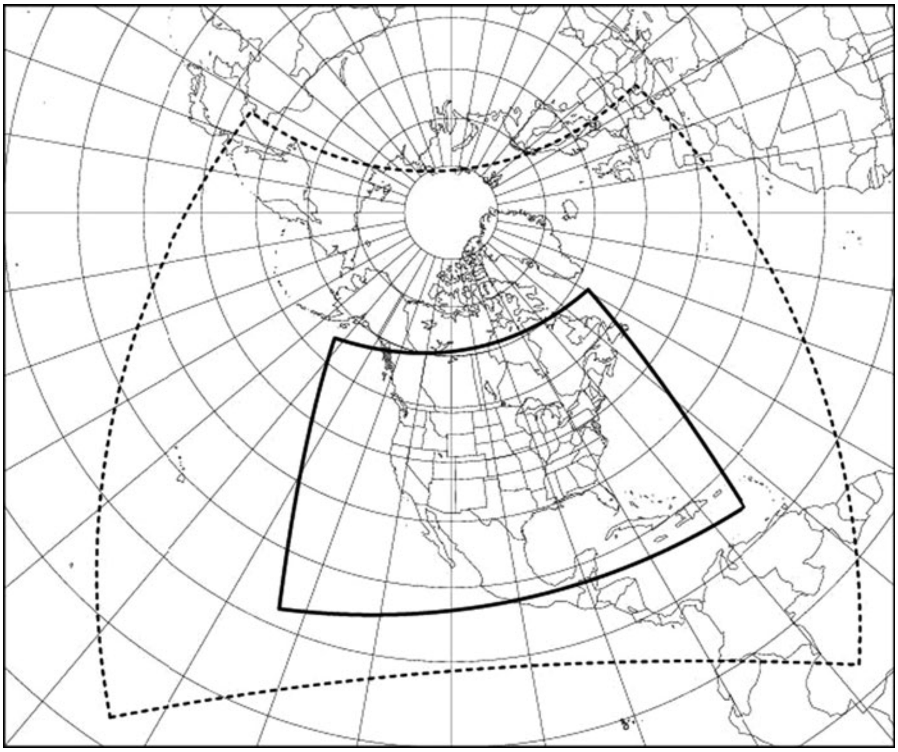

Figure 1
Model domain for NCEP regional models: *dash line* is the running domain and *solid line* is the output domain, where the low visibility and fog forecasts were evaluated

usually not reliable enough to represent fog; therefore, models tend to seriously miss the fog forecast and Vis in many cases (Gultepe *et al*., 2006; Gultepe and Milbrandt, 2007). To better detect fog, other variables besides LWC should be considered to enhance the hit rate of fog forecasting.

For this reason, the rule-based approach used in the fog prediction is suggested for the post-processor of the NCEP models (1 for yes and 0 for no fog occurrence) given as

$$\text{LWC at model lowest level} \geq 0.015\ \text{g kg}^{-1},\ \textit{OR} \tag{1a}$$

$$\text{Cloud top height (Zt)} \leq 400\ \text{m AND Cloud base height (Zb)} \leq 50\ \text{m,OR} \tag{1b}$$

$$\text{10m - Wind Speed}(U_{10m}) \leq 1\ \text{ms}^{-1}\text{AND 2 m-RH } (RH_{2m}) \geq 95\,\% \tag{1c}$$

This diagnosis is similar to the conceptual scheme suggested by Croft (1997). The LWC rule in (1a) comes from the definition of fog visibility range. With Kunkel's equation (Kunkel, 1984), LWC $\geq$0.015 g kg^{-1} is equivalent to visibility less than 1,000 m, that is defined as a threshold for fog definition by the World Meteorology Organization (WMO). The Zt threshold in (1b) follows the general features of fog. Observations indicate that the depth of most fogs on land is about 100–200 m (radiation fog). Some marine fogs or advection fogs are deeper, but rarely exceed 400 m. The Zb threshold in Eq. 1b reflects the lowest level height of our models. To deal with ground fog, the RH-wind rule (1c) is included. The selection of thresholds for surface wind and *RH* over large domains in a model is more difficult than those of LWC and cloud heights because fog usually occurs locally and different models can have different *RH* and wind biases. In many cases, fog was reported while the model RH was less than 100%. Thus, weak turbulence is usually a necessary condition for radiation fog formation. With appropriate thresholds for RH and turbulence intensity (e.g. those suggested

Table 1

Configuration of NAM, RUC and the SREF system

Models (members: ctr, p, n control, positive and negative perturbs)	Convection scheme	Micro-physics	Res (km)/levels	PBL	IC/BC	Long wave	Short wave
NAM	BMJ	Ferrier	12/60	MYJ	GDAS/GDAS	GFDL	GFDL
RUC	GD	Thompson	13/50	MYJ	GDAS/GDAS	RRTM	Dudhia
Eta (3: ctr, p1, n1)	BMJ	Ferrier	32/60	MYJ	NDAS/GENS	GFDL	GFDL
Eta (3: ctr, p1, n1)	KF	Ferrier	32/60	MYJ	NDAS/GENS	GFDL	GFDL
NMM (5: ctr, p1, n1, p2, n2)	BMJ	Ferrier	32/52	MYJ	GDAS/GENS	GFDL	GFDL
ARW (5: ctr, p1,n1, p2, n2)	KF	Ferrier	35/36	YSU	GDAS/GENS	RRTM	Dudhia
RSM (3: ctr, p1, n1)	SAS	ZC	32/28	MRF	GDAS/GENS	RRTM	NASA
RSM (2: p2, n2)	RAS	ZC	32/28	MRF	GDAS/GENS	RRTM	NASA

KF stands for the Kain-Fritsch scheme (Kain and Fritsch, 1990), SAS for the simplified Arakawa-Schubert convection scheme (Kanamitsu *et al.*, 2002), RAS for the relaxed Arakawa-Schubert (RAS) convective scheme (Kanamitsu, *et al.*, 2002), Ferrier for the Ferrier mircophysical scheme (Ferrier, 2002), ZC for the Zhao and Carr micro-physics scheme (Zhao and Carr, 1997), YSU for the Yonsei University PBL scheme (Hong and Dudhia, 2003), GDAS and NDAS for the Global Data Assimilation System, and NAM Data Assimilation System, MRF for Medium Range Forecast system (belong to NCEP global forecast system GFS) and GENS for the NCEP's Global Ensemble System (Toth and Kalnay, 1993). GD stands for Grell-Devenyi convective scheme (Grell and Devenyi, 2002), Thompson for the Thompson micro-physics scheme (Thompson *et al.*, 2004), BMJ for the Betts-Miller-Janjić convective scheme (BMJ, Janjić, 1996), MYJ for the Mellor-Yamada-Janjić scheme (Janjić, 1996), GFDL for the Geophysical Fluid Dynamics Lab schemes (long wave: Schwarzkopf and Fels, 1991, short wave: Lacis and Hansen, 1974), RRTM for the Rapid Radiative Transfer model (Mlawer *et al.*, 1997), Dudhia for the Dudhia short wave scheme (Dudhia, 1989)

for radiation fog by Zhou and Ferrier, 2008), fog in a model grid area can be diagnosed more efficiently. Unfortunately, the turbulence intensity was not an output from these models. An alternative approach is using a parameter related to a combination of surface RH and wind speed but no quantitative relationship between wind speed and turbulence intensity has been developed for fog formation. The forecasters usually use RH at 2 m ($\geq$90–100%, Gultepe *et al.*, 2007; some use dew point temperature) and 10 m wind speed ($\leq$1–2 ms^{-1}) to check for fog occurrence, and these thresholds vary depending on location and model type. For centralized fog forecasting, the optimized *RH* and wind speed thresholds as 95% and 1 ms^{-1}, respectively, were usually used by local forecasters (Justin Arnott, NWS Binghamton, NY, personal communication). This application is slightly different than that of B08RDP due to a wet bias in current NCEP regional models.

2.4. *Configuration of the NCEP SREF System*

The NCEP SREF system has four base models, including 32 km Eta (Black, 1994), 32 km WRF-NMM (or NCEP-WRF, Janjić, 2001), 32 km WRF-ARW (or NCAR-WRF, Skamarock, 2005) and 45 km RSM, the regional spectral model (Juang, 1997). Each base model is expanded with perturbed ICs to generate more ensemble members by using one control IC plus one or two pairs of positive and negative perturbed ICs (called "breeding", Toth and Kalnay, 1993) for each of the base models. These bred ensemble members are further combined with various physical parameterizations, PBL, surface layer and radiative schemes to construct a total of 21 members in the current SREF system. The configuration of different models with different parameterizations and schemes in the SREF is also listed in Table 1, which shows that the four base models are perturbed into six, five, five, and five members, respectively. All models in Table 1 use the same Noah Land Surface Model (Noah, Ek *et al.,* 2003).

Such a combination of various models with different physical schemes and parameterization, and using a variety of ICs, the simulation errors for fog forecasting can appropriately be addressed. However, for an ensemble forecast system, the total number of members, or ensemble size, is often a concern for a particular weather's ensemble probabilistic prediction. The question is, "are the 21 members in the SREF large enough to increase the forecast skill in fog prediction?" The answer can be yes because of saturation of ensemble members. According to the theory of Richardson (2001), the optimum ensemble size of an ensemble forecast system will eventually saturate. That is, an initial increase in the ensemble size has a bigger effect on the prediction skill enhancement, but if there is a further increase in the model populations, results will reach to a saturation level. At that point, there will be no more improvement in the prediction. The theory also indicates that the most effective ensemble size for a rare-occurrence weather range is within 10 members, whereas the most effective ensemble size for frequent weather events can be reduced to as low as within five members. Therefore, ensemble size of 21 for the SREF is good enough to satisfy our objectives in a fog probabilistic forecast.

2.5. *Probabilistic Distribution Function Computation*

In this study, fog types are not considered and the ensemble member models used are not tested for specific fog type. Thus, at the present stage, equal capability of the each member used in ensemble runs to capture fog is assumed. This means that each ensemble member in the SREF system has an equal-member weight. For such an equal member weight ensemble system, the ensemble probabilistic prediction, or probability distribution function (PDF), for a Vis range smaller than a threshold value ($\mathrm{Vis_t}$) for each grid number *(i, j)* at a forecast time is given as :

$$P_{i,j}(\mathrm{t,vt}) = \frac{1}{N}\sum_{m=1}^{N} K_{i,j}^{m}(\mathrm{t,vt}) \text{ for Vis} \leq \mathrm{Vis_t} \quad \text{(or fog is predicted)} \qquad (2)$$

where $K_{i,j}^{m}(\mathrm{t,vt}) = 1$ if a member *(m)* predicts $\mathrm{Vis} \leq \mathrm{Vis_t}$ in a grid at forecast time *t*. In simulations, the $\mathrm{Vis_t}$ is set up as 500, 1,000, 2,000, 4,000, 8,000 m. The *N* is the ensemble size and taken as 21 in the SREF. The $P_{i,j}$ *(t, Vis_t)* is the ensemble probability at *t* for $\mathrm{Vis} \leq \mathrm{Vis_t}$ for grid $i = 1,2, \ldots \mathrm{nx}$, and $j = 1,2, \ldots, \mathrm{ny}$, where nx and ny are max size of the model area. For example, in a grid, if there are 10 members that predict fog, then the ensemble fog prediction probability is 10/21 (47.6%). Thus, for each forecast run, the ensemble probability distribution function for various visibility thresholds can be computed over the entire domain based on Eq. (2) and validated grid to grid against observations at all grids within the domain.

3. *Validation Data and Evaluation Method*

3.1. *Validation Data*

An evaluation of fog prediction over a large domain like North America is generally difficult due to a lack of direct fog observations and the fact that model-based fog value represents a grid area that cannot be interpolated to the location of the observational sites. Thus, using high resolution visibility analysis (also gridded data) from the Aviation Digital Database Service (ADDS) of the Aviation Weather Center, NCEP, as validation truth for our objective verification is appropriate. The grid space for the ADDS data is about 5 km which is routinely analyzed from more than 5,000 surface station observations over the US and Canada through a data assimilation system. The 5 km grid space is much smaller than that of the regional models. To objectively compare the grid-scaled visibility values or fog events from the regional models against the ADDS data at the same locations, the visibility/fog forecast from each model was first downscaled to match the ADDS grid values using *copyg* (the NCEP's grid converter; Zhou *et al.*, 2011) with the nearest neighbor option (no interpolation is performed because fog is considered a non-continuous feature in the horizontal direction).

3.2. *Evaluation Method*

The observational data period covers 6 months from Nov 1, 2009 to Apr 30, 2010. This time period is chosen because of an observed high occurrence of fog events in this period. If the observed/forecast visibility is ≤1 km in a grid, the ADDS/model grid is considered as foggy. The model forecast visibility is compared to the observed visibility in a grid as follows: if visibility is ≤1 km in both observation and model grids, this is assigned as a "hit"; if forecast visibility is ≤1 km but observed is >1 km, this is assigned as a "false alarm"; and if forecast visibility is >1 km but observed is ≤1 km, the result is assigned as a "missed alarm". Using these statistical classifications, forecast scores such as bias, probability of detection (POD), and equitable threat score (ETS) can be derived. The bias here is defined by the ratio of total forecast events divided by total observed events. If the bias is larger/smaller than 1, it means the model is over/under predicting. An over-prediction system means higher false alarms but not necessarily higher hits. In comparison to the usual threat score (TS or critical success index, CSI), the ETS has an advantage that removes the random hit contribution from the score. These traditional scores can generally be used to evaluate both a single model (deterministic) forecast and an ensemble probabilistic forecast in deterministic aspect. Since ETS is an overall score that considers combined effects (POD, false alarm rate, and missing rate, etc.), the performance ranking of evaluated models will be based on the values of ETS in the latter sections.

The traditional (deterministic) scores are usually not enough to evaluate a probabilistic forecast from an ensemble forecast system. Some other probabilistic measures, such as Brier skill score, resolution and reliability, are also required as we did during an evaluation of the ensemble fog forecast in B08RDP (Zhou and Du, 2010). Since our purpose is to compare the fog predictions from a single and an ensemble system (not to evaluate the ensemble forecast system itself), only deterministic verification scores are evaluated. To compare a single model forecast and an ensemble probabilistic forecast in a deterministic aspect, the probabilistic visibility/fog forecast should be converted to a deterministic visibility/fog forecast with a certain probability threshold percentage. For a given percentage threshold (such as 50%), a probabilistic forecast can be viewed as a deterministic forecast in the way that an event (e.g. visibility <1,000 m) is expected to occur when the forecast probability is greater than or equal to the selected threshold. That is, if more than 10 out of 21 members in the SREF predict visibility ≤1,000 m in same grid, fog is expected in this grid by the ensemble forecast. To evaluate an ensemble forecast for a fog event over the entire PDF space, several probability thresholds such as 10, 20, 30, … 90, and 100% were selected to evaluate which ensemble probability thresholds will yield the best prediction performance.

4. Results and Discussions

4.1. Performance of Current Regional Models

In this section, first, results are presented for low visibility forecast from each regional model. The evaluation scores from NAM-12, RUC-13, and NMM-32 are illustrated in Fig. 2, from which the performance for fog range (visibility ≤1 km) can be estimated. Figure 2 shows that the general performances degrade as the visibility threshold decreases. For the visibility threshold of fog, the POD is about 25% for RUC-13, 10% for NAM-12 and only 5% for NMM-32. Since NAM is also a NMM-based regional model, it can be expected that the coarse resolution model NMM-32 has a lower hit rate (POD) or is more prone to miss the forecast than that of higher resolution (12 km) of the same model in fog prediction.

Another feature shown in Fig. 2b is that the POD for dense fog (visibility ≤0.5 km) is lower than that of shallow fog intensity (visibility >0.5 km but ≤1 km). In other words, dense fog events are more difficult to detect by these operational models in fog prediction. Figure 2a shows significant high biases for fog predictions by all three models (where bias ∼1 means no bias). A positive bias implies an over-prediction or a false alarm of fog forecast. For shallow fog, the highest bias is 3 (or 300%) for RUC-13. The bias for dense fog prediction is even larger. Such high positive biases for all models indicate that very low visibility or fog from all NCEP regional models is highly overpredicted. The low POD with high bias leads to poor general performances as indicated by ETS (Fig. 2c), where the ETS values for all three models are around 5%. These scores are similar to the single model evaluation in B08RDP. To compare the ETS values for fog prediction to those for precipitation prediction, the average precipitation forecast ETS (∼35%) from the same NCEP regional models is also marked in Fig. 2c, meaning that the ETS for fog prediction is much lower than that for

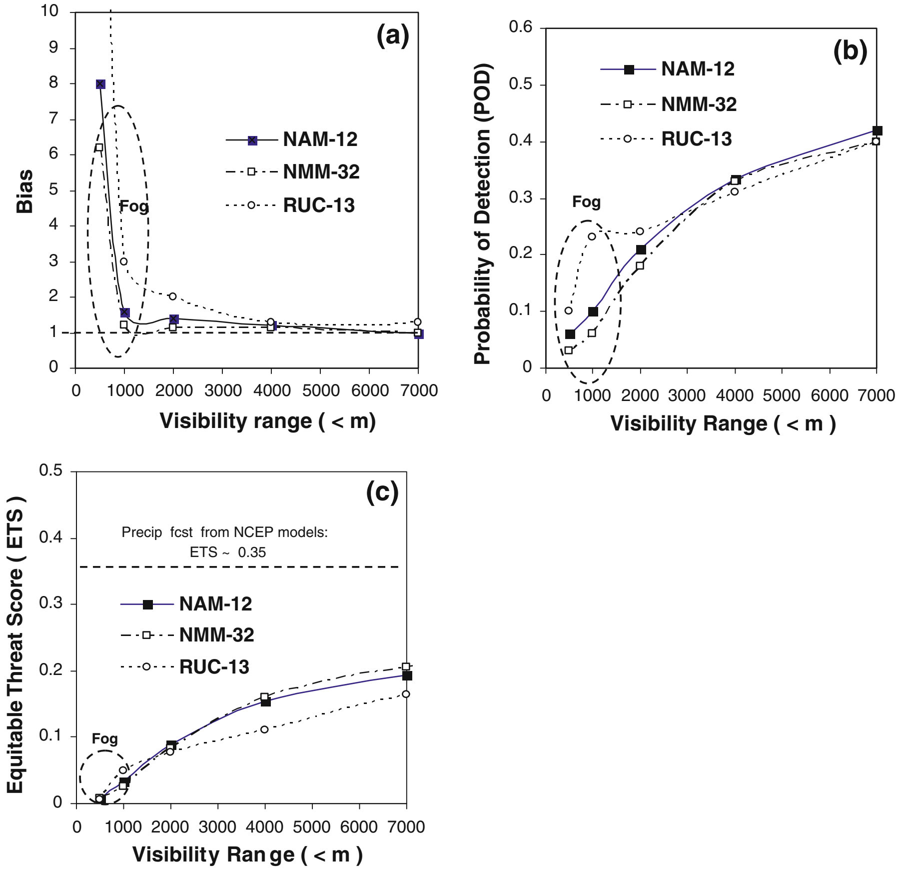

Figure 2

Tests for visibility over different thresholds (x-axis): bias (**a**), POD (**b**) and ETS (**c**) for each of the three regional models

precipitation prediction. Therefore, in order to catch up the performance of precipitation forecast at NCEP, tremendous efforts should be dedicated to improving our fog forecast. Low POD (Fig. 2b) and high bias (Fig. 2a) implies that the current models overpredict low visibility or fog occurrence in some areas but miss most of the real fog events. In fact, the visibility-diagnosed fog method is based on the LWC-rule. This is the reason why current visibility-diagnosed fog prediction, without input from other variables, has very low performance (GULTEPE *et al.*, 2006).

To examine this feature, let us further look at an east coast regional fog event that occurred on Nov 16, 2009 (Fig. 3a). This particular regional fog event covered several east coast states, including northern Florida, almost all of South Carolina and North Carolina, most of Virginia, Maryland and Delaware, extending to some regions of Pennsylvania, New York and some of Ontario and Quebec of Canada. The visibility computations in the three models are obtained from fog LWC. The green colors (dark green, green and light green) indicate the fog intensities expressed by visibility levels and locations. Comparing the observed fog location and its intensity (Fig. 3a) with the fog visibility at 12 h forecast by NAM (Fig. 3b), one can easily notice that the NAM forecast missed most of the fog events in Virginia and North Carolina although it captured some of the fog locations in Maryland and Delaware. However, it issued false alarms for half of Pennsylvania and New York states, and most of the other northeast states as well as some regions of Canada. At hour 9, the forecast by NMM-32 (Fig. 3c) almost missed the entire fog event over the east coast. This case for Vis, again, shows a worst performance of a lower resolution model than that of higher resolution model. The RUC's 12 h forecast for this case can be seen by comparing Fig. 3a and d. The RUC forecast also missed most of the fog in Maryland, Virginia, North Carolina and South Carolina, and over-predicted the fog in Pennsylvania and New York states as well as over most of the other northeast regions, similar to the NAM forecast. This case clearly illustrates the "large false alarm" feature of low visibility and fog forecast from current models and reminds us that incorrectly predicted location and amount of grid-scaled fog LWC at the surface makes it difficult to precisely compute the visibility in the case of fog.

4.2. *Suggested Improvements*

Three approaches have been directed at improving the performance of the low visibility and fog forecasting in NCEP. The first is applying the rule-based fog diagnostic scheme in the three regional models, the second is conducting an ensemble fog prediction system in SREF, and the third is a combination of the rule-based and ensemble technique for fog prediction. These are explained as below.

4.2.1 *The Rule Based Technique*

The rule-based fog diagnostic scheme has been extensively evaluated in B08RDP in China (Zhou and Du 2010). Because fog is extremely sensitive to surface variables, particularly to RH and wind speeds, selections of different threshold values in the rule will have significant impacts on the performance of the rule-based fog forecast. The sensitivity test of the RH-wind rule in B08RDP has shown that if the RH threshold is too large ($\sim$100%) or the wind threshold too small, the performance will hit a limit after which the RH-wind rule no longer has an effect on fog forecast and only the cloud rule (1b) and LWC rule (1a) play roles under such circumstances. On the other hand, if the threshold for RH is too low or the wind is too strong, the overall performance score will be even lower than that of the visibility-diagnosed method due to too many false alarms. In other words, inappropriate RH-wind thresholds may cause a negative contribution to the forecast score. Therefore, RH and wind thresholds are critical but their appropriate thresholds are more important to a successful fog forecast. The evaluation in B08RDP also revealed that with a rule-based fog detection scheme, the prediction ETS was tripled in comparison to that with the visibility-diagnosed method, in which the RH-wind rule has most of the contribution (as large as 50%) to the skill improvement. This implies that radiation fog is the most frequent fog type since RH and calm air are two critical conditions for radiation fog. One can expect that without the RH-wind rule the models would miss at least 50% of the fog events. In this study, the rule-based fog detection scheme was further tested in each of the three regional models over North America and evaluated with the same ADDS visibility analysis data. The evaluated scores for various models are listed in Table 2, in which NAM-12 shows better POD than the other two models. Despite its higher bias, NAM-12 has best overall performance indicated by the ETS. Comparing NAM-12 and NMM-32, it is demonstrated again that the forecast skill of a higher resolution model is better than that of a lower resolution peer model for fog prediction with the rule-based detection.

4.2.2 *Ensemble Fog Forecasting Technique*

This technique involves the computation of the low visibility ($\leq$1,000 and $\leq$500 m, respectively) based on ensemble predictions from the SREF system. Computing the ensemble probability for low visibility in a grid from the SREF is relatively simple: the first step is counting how many ensemble members predict low visibility in this grid, and then dividing the count by the ensemble size, 21, to obtain the probability of low visibility in this grid with Eq. 2. To use the traditional measures in evaluation of an ensemble forecast, the SREF low visibility probabilistic forecast was first converted to a deterministic forecast with a certain probability threshold. To evaluate with which ensemble forecast probability threshold the SREF has the best low visibility prediction performance, multiple forecast probability thresholds, generally from 10 to 100%, with every 10% as an interval, were selected and evaluated respectively for both visibility $\leq$1,000 and 500 m forecasts (see Fig. 4). The results reveal that (1) the

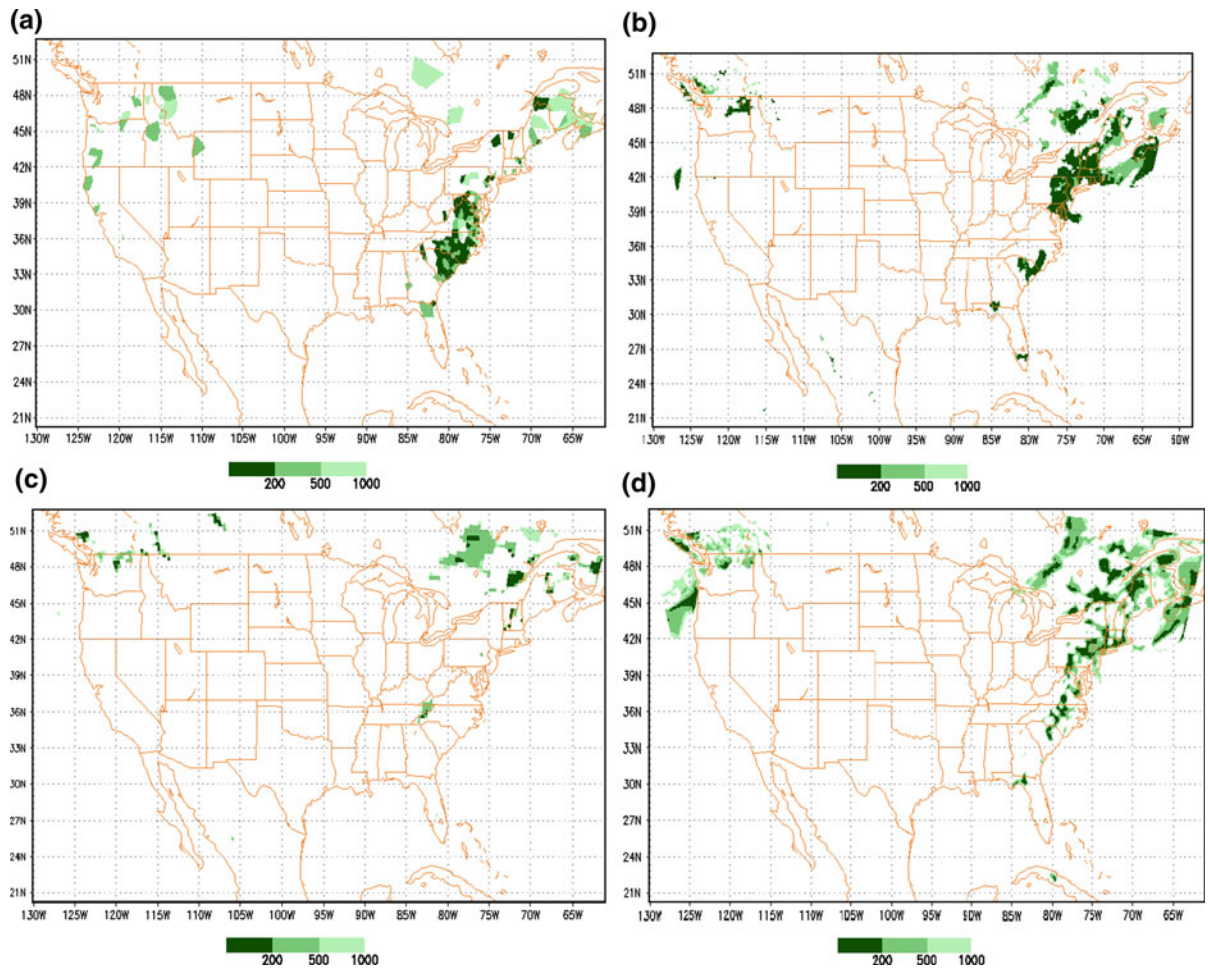

Figure 3

Nov 16, 2009 fog visibility observations from ADDS at 1200 UTC in east coast (**a**) and their 12 h forecasts from NMM-12 (**b**), 9 h forecast from NMM-32 (**c**) and 12 h forecast from RUC-13 (**d**). *Dark green* is for visibility <0.2 km, green for <0.5 km and *light green* for <1.0 km to represent different fog intensities

Table 2

Scores for rule-based fog detection method used in single models

	NAM-12	RUC-13	NMM-32
Bias	2.40	2.25	1.60
POD	0.290	0.240	0.185
ETS	0.071	0.065	0.050

performance for visibility ≤500 m forecast is consistently lower than that for visibility ≤1,000 m forecast over all of ensemble forecast probability thresholds, which means that dense fog is also more difficult to predict with an ensemble forecast system as with a single model; and (2) for different ensemble forecast probability thresholds the SREF for both low visibility ranges (1,000 and 500 m) have different forecast performances. For a smaller forecast probability threshold, the ensemble gives a higher POD (Fig. 4b) but with a large bias as a penalty (Fig. 4a). To decrease the bias, a larger forecast probability threshold should be chosen. In this case, the forecast POD decreases accordingly. Therefore, how to choose an appropriate forecast probability threshold in fog prediction means a trade-off bias and POD. Different users may select different forecast probability thresholds based on their own unique requirements, objectives, economic values (cost-loss analysis), and decision making procedures. For example, if the cost of protection is not so high in comparison to the loss, users may prefer a higher POD and may not worry about a false alarm, while others may be the opposite. If one is more concerned about POD, select a smaller forecast probability threshold; otherwise, select a larger forecast probability threshold to reduce false alarms and bias. One of the advantages of an ensemble forecast system is that it provides different users with different choices and decision making procedures based on their own needs but a single model forecast can not. Such a distribution of evaluation scores over different probability thresholds from an ensemble forecast system provides users with a decision making reference. If there is no preference, a medium range of forecast probability threshold can be selected around 40–50%, where the ensemble forecast usually has a best performance as shown in Fig. 4c. It should be noted that such a 40–50% probability range is a common feature for all of probabilistic forecast systems (Wilks, 2006).

4.2.3 Integrated Technique

This technique is a combination of the rule-based fog detection into the SREF system. The method is as follows: first apply the rule-based fog detection in each of the ensemble members from the SREF to determine whether this member predicts fog in a grid, and then use this to compute the ensemble PDF for fog occurrence with Eq. 2, based on how many ensemble members have fog occurrence in the same grid. To determine if one issues a fog forecast in a grid depends on what probability threshold is chosen. To evaluate which probability threshold in the third effort (i.e. combination of rule-based diagnosis and ensemble) has the best fog prediction performance, different probability thresholds were tested and shown in Table 3. One can see that comparing to the low visibility ensemble prediction, the fog ensemble prediction combined with the rule-based fog detection has a similar distribution of score over different forecast probability thresholds (comparing Table 3 and Fig. 4): both POD and bias (Table 3, row 2 and row 3) consistently decrease as the forecast probability threshold increases. Particularly, the ETS score (Table 3, row 4) has its best value near 40%. If choosing a smaller probability threshold, the bias will be very high although it can raise the POD. To reduce the bias or false alarms, a larger probability threshold should be used. To see how this works in an actual ensemble fog forecast, let us see the SREF fog prediction for the same case in Fig. 3. Figure 5 shows the 9 h forecast of fog ensemble PDF from the SREF over North America valid at 12Z, Nov 16, 2009. The regions where fog most likely occurred are marked with cyan-orange-red colors. Comparing the observation in Fig. 3a and the PDF forecast in Fig. 5, it can be seen that fog events on the east coast are covered by yellow–red colors (ensemble probability larger than 70–100%), in North Carolina by cyan-yellow colors (larger than 50–70%) and in South Carolina by cyan color (larger than 40–50%), significantly improving the fog predictability in comparison to the single models as shown in Fig. 3b, c, and d. Having a closer look at Fig. 5, It can be noticed that many regions are colored with low PDF (10–20%). If selecting a higher probability threshold value, e.g. 40% (cyan color in Fig. 5), the false alarm regions with small PDF (10–20%) can be filtered out, leading to better agreement with observations (Fig. 4a) and improving the ensemble forecast performance.

4.3. Comparison of the Three Techniques

The comparisons of bias, POD and ETS among the three techniques are summarized in Fig. 6. As a reference, the scores of visibility method with single mode and ensemble are also indicated. For the first technique with rule-based fog detection scheme applied in the each model, although a small bias is added in comparison to the visibility method (black bars compared to grey bars for NAM-12, RUC-13 and NMM-32 in Fig. 6a), more POD and much bigger ETS are rewarded (black bars compared to grey bars for NAM-12, RUC-13 and NMM-32 in Fig. 6b, c), increased by almost 100% of ETS scores for NAM-12 and NMM-32, 30% for RUC-13. The reason for better performance with the rule-based fog detection is that fog has various types and each type of fog has its particular formation and development mechanism. The visibility-diagnosed forecast from current regional operational models at NCEP is based on the LWC rule, which may not efficiently capture

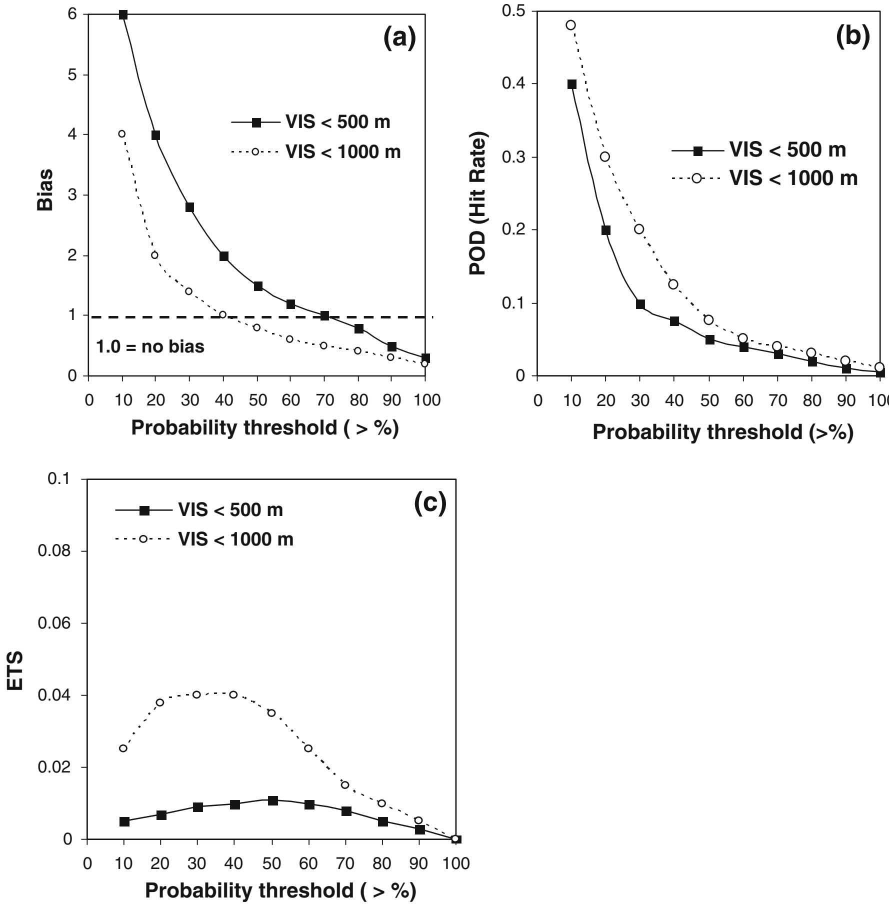

Figure 4
Scores for low visibility ensemble probabilistic prediction from the SREF: Bias (**a**), POD (**b**) and ETS (**c**) under different forecast probability thresholds as in *x*-axis

Table 3

Scores for fog probabilistic prediction for different probability thresholds from the SREF combined with the rule based fog detection method

	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
Bias	12.0	4.5	2.3	1.3	1.1	0.9	0.7	0.5	0.3	0.2
POD	0.62	0.49	0.40	0.25	0.22	0.15	0.12	0.10	0.05	0.02
ETS	0.03	0.04	0.05	0.06	0.05	0.04	0.04	0.03	0.02	0.01

all types of fog. For local fog or radiation fog, it more locally forms and develops, and in most situations, is grid-scaled weather which may not be adequately represented by the cloud schemes employed in the operational models. On the other hand, any operational model presents certain degrees of model bias, particularly, in the surface humidity, temperature and wind speed forecasts. Such biases lead to miss or false prediction of grid-scaled fog in the models in many situations and reduce the forecast POD and overall performance as a result.

Since NMM-32 is one of the base models in the SREF system, it is possible to compare the performances of the visibility-diagnosed fog detection ($\leq$1,000 m) between the SREF and the single model NMM-32 forecasts at same resolution (32 km) for the

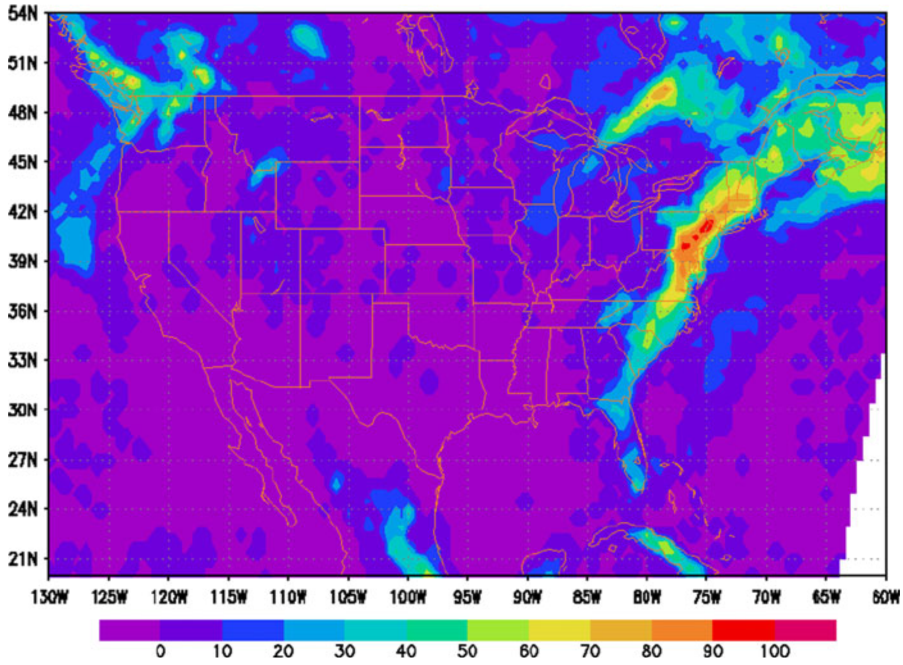

Figure 5
9 h fog ensemble probability forecast from the SREF issued at 03Z, Nov 16, 2009, valid at 12Z on the same day. The color bar is the ensemble probability indicator

second effort. Although bias reduction is not so significant for the SREF from its base model NMM-32 (compare light grey bars for NMM-32 and SREF-32 in Fig. 6a), the increases in POD (compare light grey bars for NMM-32 and SREF-32 in Fig. 6b) and ETS (compare light grey bars for NMM-32 and SREF-32 in Fig. 6c) in the SREF forecast for low visibility are obvious in comparison to those of NMM-32.

With further combination of rule-based fog detection into the ensemble in the third effort, extra POD and ETS scores were added (black bars compared to grey bars for SREF-32 in Fig. 6b and c), although a bias is expected (black bar compared to grey bar for SREF-32 in Fig. 6a). This demonstrated the better performance of ensemble over single model fog prediction over North America. It is of interest to observe that the overall score ETS of the SREF prediction, particularly with rule-based fog detection, still can not beat the same scores for NAM-12 and RUC-13 (black bar for SREF-32 compared to blacks for NAM-12 and RUC-13 in Fig. 6c). The results (ETS) of the ensemble (10 members with 15 km resolution) in B08RDP are found to be better than what is obtained from the SREF presented in this paper. It should be kept in mind that the resolution of the current SREF is 32 km, which is much lower than that of NAM-12, RUC-13 and the ensemble in B08RDP. This implies that an increase in the resolution of the ensemble system for fog prediction is an effective way to further raise its performance after the ensemble size has reached a saturation size. The horizontal resolution of the current SREF is still not high enough to skillfully predict local grid-scaled fog events even with a better fog detection scheme. This once again prompts us to increase the horizontal resolution to get a better performance for fog ensemble prediction from the SREF in the near future.

To demonstrate meteorologically why an ensemble forecast works better than a single forecast (in same model resolution), two aspects need to be explained. First, fog is a threshold weather event that is extremely sensitive to model ICs, which in general can have some errors. Small errors in the ICs will lead to totally different fog forecasts. After the ICs are perturbed around their control values in an ensemble system, the forecast can effectively

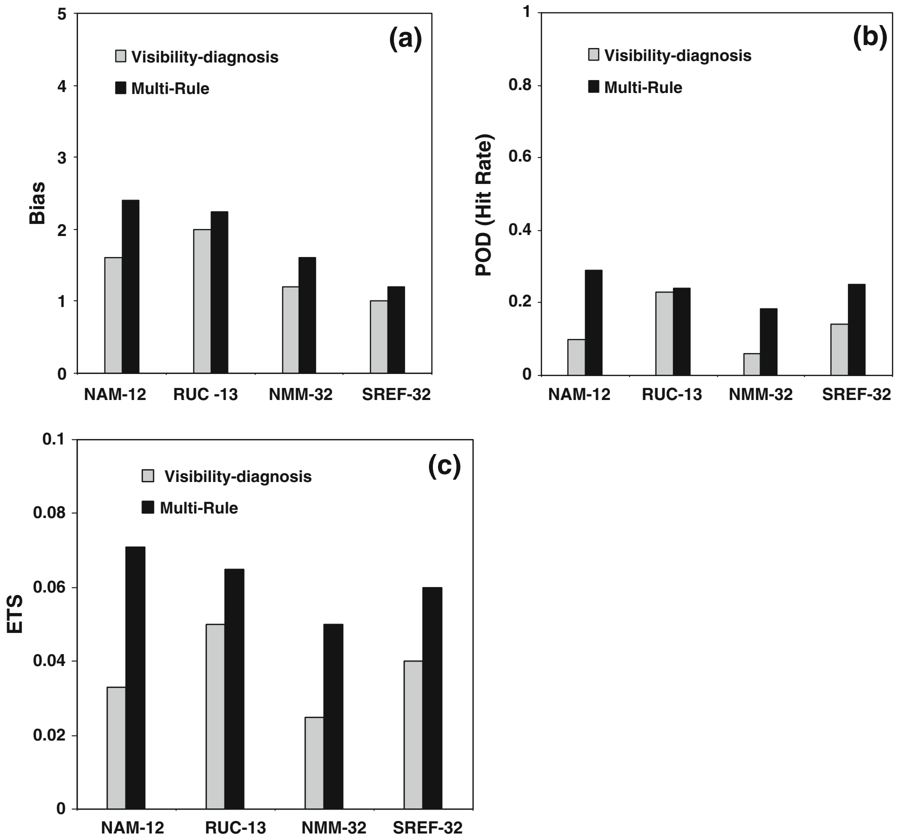

Figure 6

Bias (**a**), POD (**b**) and ETS (**c**) of fog prediction from NMM-12, RUC-13, NMM-32 and SREF with multi-rule fog detection and visibility-diagnosis schemes

encompass all possible IC values that fog may meet in the forecast. Thus, the chance of correctly forecasting fog can be significantly increased. Second, fog has various types but one model or one scheme employed may not deal with all fog types. In many cases, a model performs well for a specific fog type (Shi *et al.,* 2010) but it may not work well at all the times and over all the locations. Therefore, it is suggested an ensemble forecasting can do a better job for fog forecasting compared to the use of single model based predictions.

5. Conclusion

The operational forecasts from NCEP's three regional models, NAM-12, RUC-13 and WRF-NMM-32, over North America were evaluated against the ADDS data (observations) from November 2009 to April 2010, and their performances of low visibility and fog prediction were estimated.

The results show that the performances of the fog prediction from current models still need significant improvements. The reason may be that these models are unable to predict correct locations and intensities of fog events due probably too-coarse model resolutions, missing appropriate fog physics in the models (Gultepe and Milbrandt, 2007), and model numerical bias.

In order to improve the low visibility and fog prediction to meet the new request of NextGen of the FAA, three efforts have been made at NCEP; (a) develop an application of a rule- based fog detection scheme, (b) develop an application of multi-model and multi-physics SREF system, and (c) integrate these two applications. The rule-based fog detection includes LWC, cloud and RH-wind parameters to enhance fog detection. The ensemble application is used to address the errors and uncertainties in initial conditions, model systems, and physical schemes, and it is believed that fog is extremely sensitive to these conditions.

The validations suggested the following conclusions

- The rule-based fog detection scheme applied in the regional models doubled their forecast skill scores in comparison to visibility-diagnosed forecast from the same models.
- Ensemble fog prediction from the SREF also enhanced the prediction performance even if only with visibility-diagnosed fog detection in the SREF models. The reason is that the ensemble system can effectively encompass the perturbed initial conditions and capture various fog types with multi-models and multi-physics schemes.
- Combining rule-based fog detection into the ensemble prediction from the SREF, extra score was added to the forecast. The evaluation also indicated that if the ensemble size has been large enough, an increase in its resolution is one of critical and effective way to further raise the performance of ensemble fog prediction.

In the future, observations collected during an ice fog project (GULTEPE, *et al.,* 2008) will be tested for model performances in the cold climates. Although rule-based scheme improves the performance of fog prediction, it only predict occurrence of fog, no fog intensity can be diagnosed. In overcome this drawback, some new technique based on ZHOU and FERRIER (2008) has been suggested (ZHOU, 2011). The next step is testing and evaluation of the new scheme with both single model and the ensemble system at NCEP.

Acknowledgments

This research is, in part, in response to requirements and funding by the Federal Aviation Administration (FAA). The views expressed are those of the authors and do not necessarily represent the official policy or position of the FAA. Our special appreciation is given to the Aviation Weather Center for providing the Aviation Digital Data Service (ADDS) data to support this study.

REFERENCES

BENJAMIN, S.G., SMIRNOVA, T.G., BRUNDAGE, K.J., WEYGANDT, S.S., DEVENYI, D., SCHWARTZ, B.E. SMITH, T.L. (2003), *Application of the rapid update cycle at 10-13 km—initial testing.* Preprints, 16th Conference on Numerical and Weather Prediction, Amer. Meteor. Soc Jan. Seattle. WA.

BERGOT, T. and GUEDALIA, D. (1994), *Numerical forecasting of radiation fog. Part I: Numerical model and sensitivity tests*, Mon. Wea. Rev. *122*, 1218–1230.

BERGOT, T, CARRER, D., NOILHAN, J. and BOUGEAULT, P. (2005), *Improved site-specific numerical prediction of fog and low clouds. A feasibility study,* Wea. Forecasting *20*, 627–646.

BLACK, T.L., (1994), *The new NMC mesoscale Eta Model: description and forecast examples*, Wea. Forecasting *9*, 265–278.

CROFT, P.J., PFOST, R.L., MEDLIN, J. M., and JOHNSON, G.A. (1997), *Fog forecasting for the southern region: a conceptual model approach*. Wea. Forecasting, *12*, 545–556.

DU, J., MCQUEEN, J., DIMEGO, G., TOTH, Z., JOVIC, D., ZHOU, B. and CHUANG, H. (2006), *New dimension of NCEP SREF system: inclusion of WRF members.* Report to WMO Export Team Meeting on Ensemble Prediction System, Exeter, UK, Feb. 6–10, 2006.

DUDHIA, J., (1989), *Numerical study of convection observed during the winter monsoon experiment using a mesoscale two-dimensional model,* J. Atmos. Sci. *46*, 3077–3107.

EK, M.B., MITCHELL, K.E., LIN, Y., ROGERS, E., GRUNMANN, P., KOREN, V., GAYNO, G., and TARPLEY, J.D. (2003), *Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model,* J. Geophys. Res. 108 (D22), 8851.

FERRIER B.S., (2002), *A new grid-scale cloud and precipitation scheme in the NCEP Eta model.* Technical report, Spring Colloquium on the Physics of Weather and Climate: Regional weather prediction modeling and predictability.

GRELL, G.A., and DEVENYI, D. (2002), *A generalized approach to parameterizing convection combining ensemble and data assimilation techniques*, *Geophys. Res. Lett. 29* (14), Article 1693. doi:10.1029/2002GL015311.

GULTEPE, I., PEARSON, G., MILBRANDT, J.A., HANSEN, B., PLATNICK, S., TAYLOR, P., GORDON, M., OAKLEY, J.P. and COBER, S.G. (2009), *The fog remote sensing and modeling (FRAM) field project*, Bull. of Amer. Meteor. Soc. *90*, 341–359.

GULTEPE, I., and MILBRAND, J. (2007), *Microphysical observations and mesoscale model simulation of a warm fog case during FRAM project*, J. of Pure and Applied Geophy. Special issue on fog, edited by Gultepe I. *164*, 1161–1178.

GULTEPE, I., MÜLLER, M.D. and BOYBEYI, Z. (2006), *A new visibility parameterization for warm fog applications in numerical weather prediction models*, J. Appl. Meteor. Clim. *45*, 1469–1480.

GULTEPE, I., and ISAAC, G.A. (2004), *An analysis of cloud droplet number concentration (Nd) for climate studies: emphasis on constant Nd.* Q. J. Royal Met. Soc. *130*, Part A, 602, 2377–2390.

GULTEPE, I., PAWGOSKI. M., and REID, J. (2007), *Using surface data to validate a satellite based fog detection scheme*, J. of Weather and Forecasting *22*, 444–456.

GULTEPE, I., MINNIS, P., MILBRANDT, J., COBER, S.G., NGUYEN, L., FLYNN, C., and HANSEN, B. (2008), *The Fog Remote Sensing and Modeling (FRAM) field project: visibility analysis and remote sensing of fog n Remote Sensing Applications for Aviation Weather Hazard Detection and Decision Support*. Preprints, Edited by Wayne F. Feltz; John J. Murray, ISBN: 9780819473080, Proceedings of SPIE Vol. *7088* (SPIE, San Diego, CA), 204 pp.

HONG, S-Y., and DUDHIA, J. (2003), *Testing of a new non-local boundary layer vertical diffusion applications.* Paper# 17.3, *20th Conference on Weather Analysis and Forecasting/16th Conference on Numerical Weather Prediction*, Amer. Meteor. Soc. Jan. Seattle, WA.

JANJIĆ, Z.I, GERRITY J.P., Jr., and NICKOVIC, S. (2001), *An alternative approach to nonhydrostatic modeling*, Mon. Wea. Rev. *129*, 1164–1178.

JANJIĆ, Z.I. (1996), *The surface layer in the NCEP Eta model,* Reprints, *11th Conference on Numerical Weather Prediction*, Amer. Meteor. Soc. 354–355, Norfolk, VA.

JUANG, H.-M.H., HONG, S.-Y. and KANAMITSU, M. (1997), *The NCEP regional spectral model: an update*, Bulletin Amer. Metero. Soc. *78*, 2125–2143.

KAIN, J.S., and FRITSCH, J.M. (1990), *A one-dimensional entraining/detraining plume model and its application in convective parameterization*, J. Atmos. Sci. *47*, 2784–2802.

KANAMITSU, M., and Coauthors, (2002) *NCEP Dynamical* Seasonal Forecast Sy*stem 2000*. Bull. Amer. Meteor. Soc. *83*, 1019–1037.

KUNKEL, B.A., (1984), *Parameterization of droplet terminal velocity and extinction coefficient in fog models*, J. Climate Appl. Meteor. *23*, 34–41.

LACIS, A.A., and HANSEN, J.E. (1974) *A parameterization for the absorption of solar radiation in the earth's atmosphere*, J. Atmos. Sci. *31*, 118–133.

LEITH, C.E., (1974) *Theoretical skill of Monte Carlo forecasts*, Mon. Wea. Rev. *102*, 409–418.

MLAWER, E.J., TAUBMAN, S.J., BROWN, P.D., IACONO, M.J. and CLOUGH, S. A. (1997) *Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave,* J. Geophys. Res. *102*, (D14), 16,663–16,682.

RICHARDSON, D. S., (2001), *Measures of skill and value of ensemble prediction systems, the interrelationship and the effect of ensemble size*, Quart. J. Royal Meteor. Soc. *12,* 2473–2489.

ROGERS, E., Ek, M., FERRIER, B.S., GAYNO, G., LIN, Y., MITCHELL, K., PONDECA, M., PYLE, M., WONG, V.C.K., and WU, W.-S., (2005), *The NCEP North American Mesoscale Modeling System: final Eta model/analysis changes and preliminary experiments using the WRF-NMM*. Paper# 4B.5, *17th Conference on Numerical Weather Prediction*. Amer. Meteor. Soc., Washington D.C.

SCHWARZKOPF, M.D., and FELS S.B., (1991), *The simplified exchange method revisited: An accurate, rapid method for computation of infrared cooling rats and fluxes*. J. Geophys. Res. *96*, 9075–9096.

SHI, C., WANG, L., ZHANG, H. and DENG, X., (2010), *Experiments on fog prediction based on multi-model*, this issue.

SKAMAROCK, W.C., KLEMP, J.B., DUDHIA, J., GILL, D.O., BARKER, D.M., WANG, W. and POWERS, J.G. (2005), *A description of the Advanced Research WRF*, Version 2, *NCAR Technical Note*.

SOUDERS, C.G. and coauthors, (2010), *NextGen weather requirements: an update*. Preprint, 14th *Conf. on Aviation, Range, and Aerospace Meteorology*, Atlanta, GA, Amer. Meteor. Soc.

STOELINGA, T.G. and WARNER, T.T. (1999), *Nonhtdrostattic, mesobeta-scale model simulations of cloud ceiling and visibility for an east coast winter precipitation event*, J. Apply. Meteor. *38*, 385–404.

THOMPSON, G., RASMUSSEN, R.M. and MANNING, K. (2004), *Explicit forecasts of winter pre-cipitation using an improved bulk microphysics scheme. Part 1: Description and sensitivity analysis*, Mon. Wea. Rev. *132*, 519–542.

TOTH, Z. and KALNAY, E. (1993), *Ensemble forecasting in NMC: The generation of perturbations*, Bull. Amer. Meteor. Soc. *74*, 2317–2330.

WILKS, D.S., (2006), *Statistical methods in atmospheric sciences*, 2nd edn, International Geophysics Series, *Academic Press, 59*, 627.

ZHAO, Q. and CARR, F.H., (1997), *A prognostic cloud scheme for operational NWP models*, Mon. Wea. Rev. *125*, 1931–1953.

ZHOU, B. and B. FERRIER, S. (2008), *Asymptotic analysis of equilibrium in radiation fog.* J. Appl. Meteor. and Clim. *47*, 1704–1722.

ZHOU, B and DU J. (2010), *Fog prediction from a multimodel mesoscale ensemble prediction system*, Wea. Forecasting *25*, 303–322.

ZHOU, B, DU J., LIU S. and DIMEGO, G. (2011), *Verifications of simulated radar reflectivity and echo-top Forecasts at NCEP*, Paper# P.90, 24th Conf. on Weather and Forecasting, Amer. Meteor. Soc, 23–27, Jan 2011, Seattle, WA.

ZHOU, B., (2011), *Introduction to a new fog diagnostic scheme*, NCEP Office Note 466, US Department of Commerce, NOAA, NWS, NCEP, 33.

(Received November 8, 2010, revised March 17, 2011, accepted April 9, 2011)